DOCUMENT RESUME

ED 448 764 IR 020 527

AUTHOR	Wade, Vincent P.; Lyng, Mary
TITLE	An Automated Evaluation Service for Educational Courseware.
PUB DATE	2000-11-00
NOTE	7p.; In: WebNet 2000 World Conference on the WWW and Internet Proceedings (San Antonio, TX, October 30-November 4th, 2000); see IR 020 507.
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Computer Software Evaluation; Computer Uses in Education; *Courseware; Design Preferences; *Evaluation Criteria; *Evaluation Methods; Foreign Countries; Higher Education; Instructional Design; *Questionnaires; Student Surveys; World Wide Web
IDENTIFIERS	*Usability

ABSTRACT

This paper proposes an automated, third party World Wide Web (WWW)-based evaluation service (i.e., an online questionnaire) that focuses on usability issues of WWW-based courseware and can be used by any WWW course instructor/student. Requirements related to the usability of educational systems are summarized, including the learning environment, layout requirements of a learning interface, and presentation of instructional material components (i.e., textual material, non-textual material, and color). A table presents evaluation criteria and a comparison with other questionnaires. Advantages of the evaluation service over other questionnaires are described. Results are presented of employing this evaluation service twice on a subsection of an already proven WWW-based courseware system and once on an entire courseware offering. The paper concludes with an assessment of the benefits of using such an evaluation service to improve WWW-based courseware. (Contains 21 references.) (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 448 764

An automated Evaluation Service for Educational Courseware

Vincent P. Wade, Dept. of Computer Science, Trinity College, Dublin 2, Ireland
Mary Lyng, Dept of P & Q, Waterford Institute of Technology, Waterford, Ireland

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

G.H. Marks

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Abstract: The World Wide Web (WWW) has been heralded as a significant opportunity to deliver successful tele-educational courseware and learning experiences. Because of its ease of use, the WWW has caught the imagination of many groups of people including educators. However, the instructors who develop educational experiences are rarely very skilled in interface design. This paper proposes an automated, third party WWW based evaluation service which focuses on usability issues of WWW based courseware and which can be used by any WWW course instructor/student. This paper researches the design, development and trialling of a WWW based evaluation service for WWW courseware. The paper then describes the facilities implemented by such a web based evaluation service and presents the results of employing this evaluation service twice on a subsection of an already proven WWW based courseware system and then once on an entire courseware offering. The paper concludes with an assessment of the benefits of using such an evaluation service to improve WWW based courseware.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

1 Introduction

The last three years has seen a rapid increase in the use of web based educational tools, courseware and environments (Wade et al. 98). The central tenant of the WWW being that any individual or group of individuals (e.g. educators) can become a publisher of (educational) material. However, it is also well accepted that educational technology, over the last thirty years, has failed to achieve the revolutions that were initially forecast (Bates 95). There have been many technical reasons proposed for the failures of various educational systems, the two most widely reported being (i) the lack of pedagogical support within the educational systems (Maurer 97) (ii) the difficulties experienced by learners in using such systems (Turoff 95). This paper addresses this second most common cause of failure. Usability as defined by the International Standards Organisation (ISO) is the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments. Satisfaction is the comfort and acceptability of using the system (Macauley 95). The paper focuses on evaluating the levels of user satisfaction with presentation issues of educational WWW based courseware using a web based evaluation service.

2 Usability for Educational Systems

There has been a significant level of research into user interface design relevant to WWW based courseware. This section briefly summarises the important requirements related to user satisfaction for learning interfaces.

2.1 The Learning Environment

A learning environment must provide the learner with four key requirements (Duchastel 97) namely, information, interactivity, structure, and communication. The learning environment must also implement key interface design concepts (Jones et al. 95) which can be loosely categorised by research facilities, presentation details, integration across various media, appropriate use of tools, and help facilities.

2.2 Layout requirements of a Learning Interface

When planning the layout of the interface, many factors will require consideration including user engagement; the HCI principles categorised as Naturalness, User Support, Non Redundancy, Consistency and Flexibility (Macauley 95); interface structure; and the issues involved in the layout of instructional material such as document length, use of icons, and scrolling.

IR020527

BEST COPY AVAILABLE

2.3 Presentation of Instructional material components

The components that are required to produce instructional material can be a combination of text, graphics, video images, etc. so many factors require considering. Another issue of concern here is the inclusion of colour.

Textual Material: When deciding on how to present text, there are guidelines that need to be adhered to (Rambally et al. 87). These include: placing key information such as urgent messages and instructions in a prominent and consistent location; standardising the terminology used; establishing prompt-string conventions (captions) in order to provide a systematic, predictable visual correlation between prompts and their corresponding data input fields; and positioning captions in a natural and consistent physical relationship to the corresponding data fields. It has also been suggested that screens are easier to read if the text is structured in natural eye sequences, such as from top to bottom (Hazen 85). When writing textual educational material, guidelines on how to lay it out should be adhered to also (Bailey et al. 91). These include paragraph formatting, font size, use of hyphens, and when to use upper case.

Non textual material: Interactive graphing, windowing, and animation differentiate the computer as a medium from most other media (Hazen 85). Tables can be used to graphically represent complex relationships (Schlegal 96). Graphics should be used when they contribute to the understanding of the text (McFarland 95). However, considerations of the size of the graphic should be made, as it should be able to fit into the graphical browser's window in order to provide the learner with the whole picture, and note also that a large graphic can take a long time to download. The minimum goal for response times should be to get pages to users in no more than ten seconds, since that is the limit of people's ability to keep their attention focused while waiting (Nielsen 97).

Colour: Colour has some uses when presenting material (Rambally et al. 87). These include: linking logically related data; differentiating between required and optional data; highlighting errors; separating various screen areas such as prompts and commands; emphasising key points; and communicating the overall structure. When applying colour, the following guidelines have been recommended (Macauley 95): use colour coding consistent with user expectations; similar colours should denote similar meaning; avoid use of extreme colour pairs to avoid frequent refocusing and visual fatigue; adults may need higher brightness levels to distinguish colours; colour blind users must be considered, as they may not be able to distinguish some colour combinations; use background colours in large blocks; group related elements by using a common background colour; use bright colour for emphasis and weaker colours for background areas; and brightness and saturation draw attention. In relation to WBI, there have been recommendations made (Jones et al. 97). These include: selectable areas should be clearly identified by a royal blue colour; interactivity is apparent by changes in the cursor as the cursor is moved to a hot spot; when a selection is made, the HTML standard is to immediately change to a dark red selection; and to indicate progress made, accessed links become a light red colour.

3 Evaluation Criteria

The objective of the evaluation service is not to evaluate the semantics of instructional material content but to identify any problems with the current interface that hinders or reduces the learner satisfaction and hence hinders the learning that occurs. The evaluation service needs to ascertain whether the present learning environment's user interface is conducive to learning, i.e measure it's user satisfaction level in terms of the learner's experience. The approach of setting product usability goals and objectives (Mandel 97) was adopted to enable the setting of user satisfaction goals/targets. From these goals/targets, evaluation criteria were derived to measure the desired performance. The concentration of the evaluation was on subjective rather than performance measures because these determine the users' attitude with regard to how easy the interface is to learn, use and remember. The goals are shown at HREF 1, and resulting evaluation criteria are shown at HREF 2.

Assessing whether these criteria are adhered to requires gathering responses from individuals who have experience using the courseware being evaluated. Therefore it was decided that:

1. Courseware students participate in the evaluation.
2. A questionnaire should be used to gain feedback from the evaluation participants.

Each evaluation criterion was examined and rewritten as question statements followed by an interval response scale of strongly disagree, disagree, agree, and strongly agree. All the evaluation items were not explicitly included in the questionnaire as some items were combined with other items or were deemed to have been redundant. A list of thirty questions was produced. Rather than having a long list of questions it was decided to structure the questionnaire into five sections, one for each of the HCI principles as defined previously in section 2.2. Each question statement was examined and then placed into the appropriate questionnaire section. The full set of questions is shown at HREF 3.

3.1 Comparison with other Questionnaires

There are evaluation forms available on-line for measuring the usability of courseware on the Web, but they tend to cover all aspects of usability. Therefore, the level of user satisfaction with regards to presentation issues is only partially measured, as it is only one aspect of the questionnaire. Table 1 contains a comparison between four on-line questionnaires and the one presented in this paper. The questions compared are those, which are applicable to measuring the level of user satisfaction with regards to presentation issues of WWW based courseware. The questionnaires surveyed are included in HREF 4, HREF 5, HREF 6 and HREF 7.

	Toolbox	Comp 200	QUIS (OLT)	Henke	New Questionnaire
Appropriate Language Used					X
Style Used	X				X
Cluttered Screen	X	X		X	X
Natural eye movement	X	X			X
Use of extra features	X	X		X	X
Text vs Background Colour	X				X
Link Colour				X	X
Visited Link Colours				X	X
Meaning of Icons				X	X
Intuitive Navigation	X	X	X	X	X
Amount of Scrolling					X
Adequate Error Message					X
Adequate instructions	X		X		X
Availability of Instructions					X
Useful Examples	X				X
Adequate amount of Examples	X				X
Bookmarking facilities					X
Interface Perception				X	X
Feedback is prompt and useful	X	X	X	X	X
Consistent Terminology					X
Consistent Style					X
Consistent Error Message					X
Consistent Colour					X
Consistent Navigation				X	X
Consistent Headings		X			X
Amount of material per screen	X			X	X
Repetition of material					X
Repetition of navigational aids	X				X
Availability of Quit					X
Access to any module	X				X
Inclusion of Exercises	X				
Print Out facility		X			
Collaboration Facilities		X			
Overall Reaction			X		
Frame Use				X	
Search /Index facility				X	
Section Complete				X	
Provision of Cues				X	

Table 1: Comparison of on-line questionnaires for evaluating the levels of user satisfaction with the presentation of web based courseware

4 Developing an 'Evaluation Service' not just another on-line questionnaire

In order to validate the proposed questionnaire, an entire WWW based 'evaluation service' was developed. The main objective of the service was to allow any WWW course designer/instructor to register their course with the evaluation service and embed hyper-links from their course to the evaluation service. The service allowed instructors to password protect access to their own course and customise the questionnaire for their own course. By embedding a link to the evaluation service, the instructors were able to provide seamless access from their courses to the evaluation service. After the students completed the evaluation questionnaires, the instructors were able to retrieve an automatically generated detailed statistical breakdown of the evaluation results as well as a qualitative analysis and suggestions as to where the courseware could be improved. More specifically the evaluation service facilitated the following:

1. It allows *any* Course director to set up a course for evaluation. A course can be set up for evaluation by providing the evaluation service with her course's details. A password is chosen at this point for the course
2. Because a course director may wish not to include all questions, the evaluation service allows the course director to view all the questions available and deselect any questions that are deemed undesirable.
3. The on-line questionnaire is made available to the participant for completion. The participant must respond to all questions presented for the courseware being evaluated and may optionally provide comments and specify any relevant module/page titles for any of the questions.
4. It allows the course director to request the automatic summary and analysis of the responses submitted by participants for their course. The analysis produces Web pages of results for each section of the questionnaire. The analysis also produces Excel files of results available for downloading. The resultant analysis and summary not only allows the identification of specific weaknesses in the courseware but also the location of pages where this was perceived to occur.
5. It allows the course director reset the courseware when the analysis is complete; this facility deletes all generated analysis files and completed questionnaires making the course available to be evaluated by another group of individuals if desired.

5 Validating the Evaluation Service – Case Study

In order to validate the evaluation service, several iterations of using the evaluation service were performed. An existing, heavily used, WWW based self learning courseware on 'Relational Databases', was chosen as a validation study. The course, offered in Trinity College Dublin had previously been designed and implemented. Initially, a subsection of the courseware was chosen for evaluation. The modules chosen were the modules pertaining to the Select statement, which is one of SQL's data manipulation statements.

5.1 Iteration One Conclusions

The participants for iteration one and two of the case study were fourth year students on the BSc in Commercial Software Development, in Waterford Institute of Technology. Fourteen students participated in iteration one.

After reviewing the questionnaire results, some problems were highlighted, which reduced the level of user satisfaction with the courseware. The quantitative values that resulted from the responses made indicated that there was dissatisfaction with: the colours chosen for the links; the amount of scrolling required; and the repetition of menu items. The textual comments submitted included: there was a misuse of italics and inconsistent use of colours used for highlighting; on-line exercises and more examples needed to be included; it was not obvious how to proceed from a module to the roadmap menu; and feedback provided when a page is bookmarked was inadequate.

5.2 Iteration Two Conclusions

Twenty six students participated in iteration two. After reviewing the questionnaire results, it was concluded that there was a perceived improvement in the level of user satisfaction with the SQL courseware. There was an increase in satisfaction regarding: the use of colour; the reduced amount of scrolling; the elimination of menu items in the module headings and at the end of modules; the amount of examples used; and the error messages provided. However, one area that caused the participants difficulty was the facility to include Module exercises.

5.3 Iteration Three Conclusions

The changes applied to the subsection that proved successful were applied to the entire courseware. Other changes were applied also to enhance the interactivity of the courseware. The students who undertook the SQL courseware are from two-degree programmes in Trinity College, Dublin – BA MOD (i.e. pure computer science degree) and CSLL (i.e. computer science and linguistics degree). From the results and the textual comments volunteered by the participants, the following points can be concluded: the on-line exercises proved successful, but a change needed to be applied to the message displayed as a result of an error; the printable version of the notes also proved to be successful, but the images (tables etc.) that are displayed on screen needed to be available in the printable version also; 23% of participants felt that the icons were not intuitive, which is a significant level of dissatisfaction; the level of detail given for the module Oracle on Vax1 was not sufficient; more examples need to be included which introduces each new concept and statement; some of the windows that are used cover the Control Panel thus blocking out the facilities they provide; and the index and site map facility are not being utilised. The facility to cite a module/page that illustrated the response made proved useful in the analysis of the results, as participants used the text box provided to add comments also.

5.4 Conclusions on Evaluation Service results

For each question, the evaluation service provides numeric results in the form of percentage values that are generated from questionnaire responses. These percentage values identify strengths and weaknesses in the courseware interface. To provide some analysis for each question, the evaluation service may display a comment after the table of quantitative results have been output. If the number of responses that agree or strongly agree is greater than 60% then a comment from the satisfactory file is displayed, however if the number of responses that disagree or strongly disagree is greater than 60% then a comment from the unsatisfactory file is displayed. The comment for an unsatisfactory result indicates that there is a problem and that the associated feature needs attention.
At the end of each of the questionnaire sections, there is an overall percentage value output for each response value. The evaluation service may display a message indicating if the question responses for the section were very satisfactory (agree and strongly agree >70%), satisfactory (agree and strongly agree >50%), unsatisfactory (disagree and strongly disagree >50%), or very unsatisfactory (disagree and strongly disagree >70%).

The evaluation service does not perform a very high degree of statistical analysis, however an excel file for each section of the questionnaire is generated and is available for downloading.

The output generated from the response files displays the comments submitted by the participants also. These comments are output in two forms:

1. If a participant cites a module name or inputs a comment after a question, it is recorded with the question response. Therefore, when the evaluation service generates and outputs the results from the response files for each question, the module or comment submitted is output also. This can pinpoint where satisfaction/dissatisfaction with a feature resides.
2. The comment area at the end of each section of the questionnaire may be utilised by participants to submit general comments about the courseware. These comments are recorded and output by the evaluation service at the end of each result section. These comments can highlight problem areas also.

From a course instructor/designer viewpoint, the results output by the evaluation service are easy to read as each question is output followed by a table of responses, response percentage values, and the modules cited/comments given for each response value. As stated previously, the evaluation service may display a comment after the table of quantitative results for each question have been output. If there is dissatisfaction with a question, the comment output indicates that there is a problem and that the associated feature needs attention.

6 Conclusions

The use of such an evaluation service is an important component in the continuous improvement of successful WWW based courseware, and is an important tool in the rapid prototyping approach to successful courseware development. The criteria and questionnaire provide an explicit, easy to use method of evaluating users' satisfaction levels with presentation issues of educational systems. Although conceptually simple, the well-researched nature of the questionnaire and labour saving aspects of the evaluation service has led to a very successful tool to empower non-technical educators to improve their on-line courseware. In particular the labour saving features of the WWW based service as follows:

1. The evaluation service collected responses in a standardised format, thus aiding the quick production of quantitative results. The service not only performed the statistical analysis, but also provided an Excel file for further analysis/recording by the instructor.
2. It greatly reduced the overhead in processing the evaluation.
3. It provided semi automatic assessment of the study results. E.g. if 70% or more Agree or Strongly Agree with all statements for a particular HCI principle then the adherence to that principle is very satisfactory.
4. It promotes and encourages learner feedback and evaluation as it greatly reduced their effort in completing the questionnaire.

7 References

Bailey, H. J., Milheim, W. D.(1991). A comprehensive model for designing interactive video based materials. *Ninth Conference on Interactive Instruction Delivery*, 1991, Society for Applied Learning Technology Conference, Orlando, Florida.

Bates, A. W. (1995). *Technology, Open Learning and Distance Education*. London : Routeledge.

Duchastel, P. (1997). A Web-Based Model for University Instruction. *Journal of Educational Technology Systems* 25 (3), 221-228.

Jones, M. G., Okey, J. R. (1995). *Interface Design for Computer Based Learning Environments*. (http://129.7.160.78.InTRO.html

Jones, M. G., Farquhar, J. D. (1997). User Interface Design for Web Based Instruction. In B. Khan (Editor), *Web-Based Instruction*, 239-244. Educational Technology Publications.

Macauley, L. (1995). *Human-Computer Interaction for Software Designers*. International Thomson Computer Press.

Mandel, T. (1997). *The Elements of user interface design*. Wiley Publications.

Maurer, H. (1997). Necessary Ingredients of Integrated Network Based Learning Environments. *Proceedings of EdMedia 1997*. Keynote S[illegible]

McFarland, R. D. (1995). Ten Design Points for the Human Interface to Instructional Multimedia. *T.H.E. Journal*. 22 (7), 67-69.

Nielsen, J (1997). The Need for Speed. http://www.useit.com/alertbox9703a.html

Rambally, G.K., Rambally, R. S. (1987). Human Factors in CAI design. *Computers in Education*. 11 (2), 149-153.

Turoff, M. (1995). Designing a Virtual Classroom(TM). *International Conference on Computer Assisted Instruction ICCAI '95*. 1995, National Chiao Tung University Hsinchu, Taiwan.

Schlegal, L. (1996). *General Issues and Limitations*. http://www.netspot.unisa.edu.au/eduweb/Media/General/limits.htm

Wade, V., Power C. (1998). Network based Delivery of Automated Management of Virtual University Coursewares. EdMedia & EdTelecom98 World Conference on Educational Multimedia & HyperMedia and World Conference on Educational Telecommunications, AACE, 1998, Frieburg, Germany.

HREF 1: http://www.wit.ie/research/goals.htm

HREF 2: http://www.wit.ie /research/objectives.htm

HREF 3: http://www.wit.ie /research/questionnaire.htm

HREF 4: Toolbox – An online evaluation form for Web-based course supplements
http://web.syr.edu/~maeltigi/Toolbox98/sumeval2.htm

HREF 5 Comp 200 Evaluation Questionnaire http://ccism.pc.athabascau.ca/html/courses/comp200/cover/eval.htm

HREF 6: Evaluating Web-Based Instruction http://www.scis.nova.edu/~henkeh/hciproj.pdf

HREF 7: Questionnaire for User Interaction Satisfaction On-Line Tutorial (OLT) http://lap.umd.edu/q7/QUIS.html